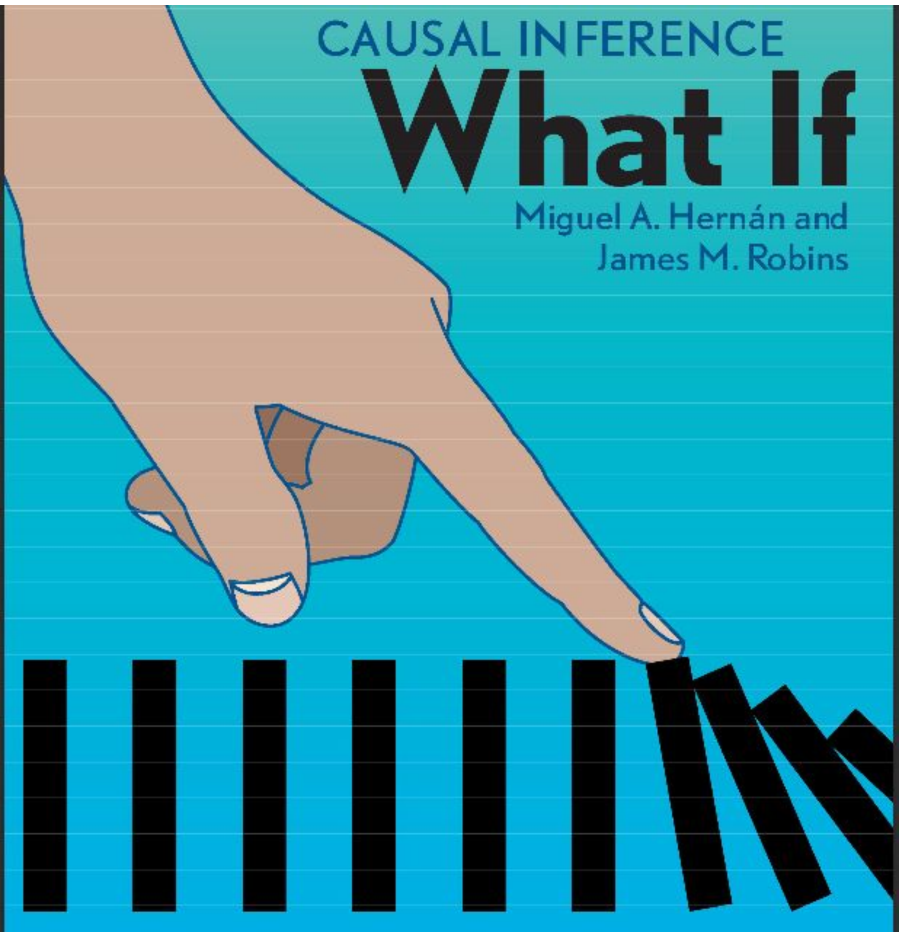
book club: chapters 8.4 – 8.6



> selection bias &
> censoring

> adjustment
> for selection bias

> selection without bias

Selection bias **review**

```
n=1000
coin1 = rbinom(n,size=1,prob=0.5)
coin2 = rbinom(n,size=1,prob=0.5)

flips <- data.frame(coin1=coin1, coin2=coin2, selected= coin1 | coin2)
head(flips)
```

	coin1 <int>	coin2 <int>	selected <lgl>
1	0	0	FALSE
2	1	1	TRUE
3	1	0	TRUE
4	1	1	TRUE
5	0	0	FALSE
6	0	0	FALSE

6 rows

Hide

```
prob.coin= subset(flips,coin2==0)

paste0("p[coin1=head|coin2=tail]: ",signif(sum(prob.coin$coin1)/nrow(prob.coin),3))
```

```
[1] "p[coin1=head|coin2=tail]: 0.498"
```

simultaneously flipping two coins

tail (0)
coin1

head (1)
coin2

Source: Causal Inference; Hernan and Robins; 2019

Selection bias **review**

```
prob.coin= subset(flips,(selected==TRUE))
head(prob.coin)
```

	coin1 <int>	coin2 <int>	selected <lgl>
2	1	1	TRUE
3	1	0	TRUE
4	1	1	TRUE
7	1	1	TRUE
8	1	1	TRUE
10	1	1	TRUE

6 rows

Hide

```
prob.coin= subset(prob.coin,coin2==0)

paste0("p[coin1=head|coin2=tail]: ",signif(sum(prob.coin$coin1)/nrow(prob.coin),3))
```

```
[1] "p[coin1=head|coin2=tail]: 1"
```

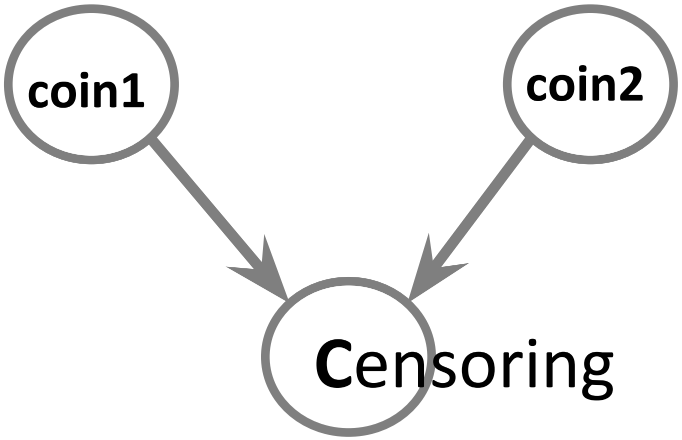
Source: Causal Inference; Hernan and Robins; 2019

Selection bias **review**

```
prob.coin= subset(flips,(selected==TRUE))
head(prob.coin)
```

	coin1 <int>	coin2 <int>	selected <lgl>
2	1	1	TRUE
3	1	0	TRUE
4	1	1	TRUE
7	1	1	TRUE
8	1	1	TRUE
10	1	1	TRUE

6 rows

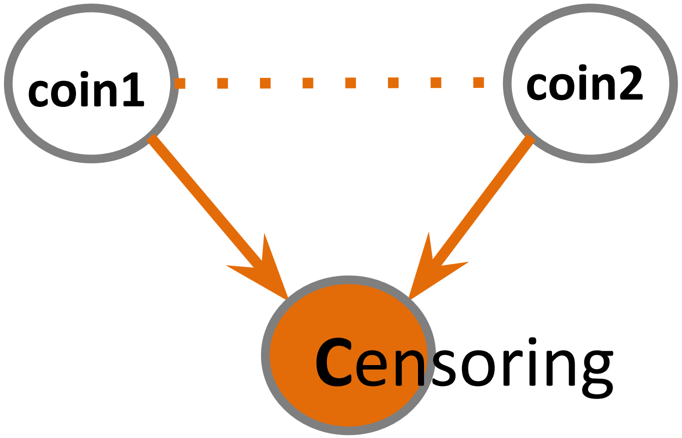


Hide

```
prob.coin= subset(prob.coin,coin2==0)

paste0("p[coin1=head|coin2=tail]: ",signif(sum(prob.coin$coin1)/nrow(prob.coin),3))
```

```
[1] "p[coin1=head|coin2=tail]: 1"
```



Source: Causal Inference; Hernan and Robins; 2019

8.4 Selection Bias and **censoring**

case I: we do get data from all participants

$$Y^a \perp\!\!\!\perp A$$

60 participants

A

*treatment (**randomized**)*:
1: wasabi (50% chance)
0: no wasabi (50% chance)

30

30

??? causal

Y

outcome after one year:
1: death
0: alive

	alive	dead	
A=0 (- wasabi)	13	17	30
A=1 (+ wasabi	13	17	30
	26	34	60

associational risk ratio:
Pr[Y=1|A=1] / Pr[Y=1|A=0] = 1
causal risk ratio:
$Pr[Y^{a=1}=1]$ / $Pr[Y^{a=0}=1]$ = 1

Source: Causal Inference; Hernan and Robins; 2019

8.4 Selection Bias and **censoring**

case I: we do get data from all participants

$Y^a \perp\!\!\!\perp A$

60 participants

A

*treatment (**randomized**)*:
1: wasabi (50% chance)
0: no wasabi (50% chance)

30
30

??? causal

Y

outcome after one year:
1: death
0: alive

U

unmeasured:
atherosclerosis

L

Yes (L=1): 40
No (L=0): 20

heart disease (L)

	alive	dead	
A=0 (- wasabi)	13	17	30
A=1 (+ wasabi	13	17	30
	26	34	60

Source: Causal Inference; Hernan and Robins; 2019

8.4 Selection Bias and **censoring**

case II: we do **NOT** get data from all participants

$Y^a \perp\!\!\!\perp A$

60 participants

A

*treatment (**randomized**)*:
1: wasabi (50% chance)
0: no wasabi (50% chance)

30

30

??? causal

U

Yes (L=1): 40
No (L=0): 20

outcome after one year:
1: death
0: alive

Y

unmeasured:
atherosclerosis

L

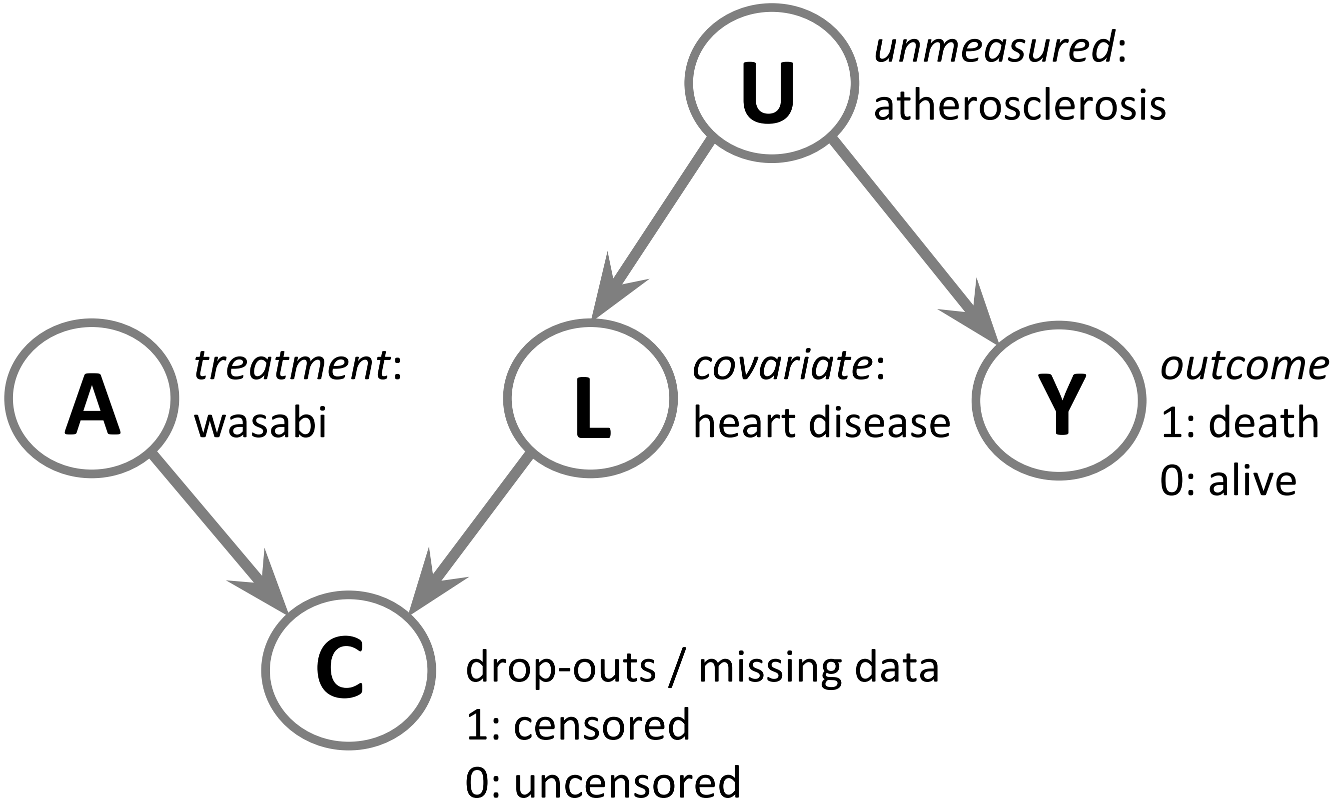
heart disease (L)

	alive	dead		follow up	observed dead
A=0 (- wasabi)	13	17		22	11
A=1 (+ wasabi	13	17		9	4
	26	34			

associational risk ratio:
(4/9) / (11/20) = **0.89**
on uncensored data

8.4 Selection Bias and **censoring**

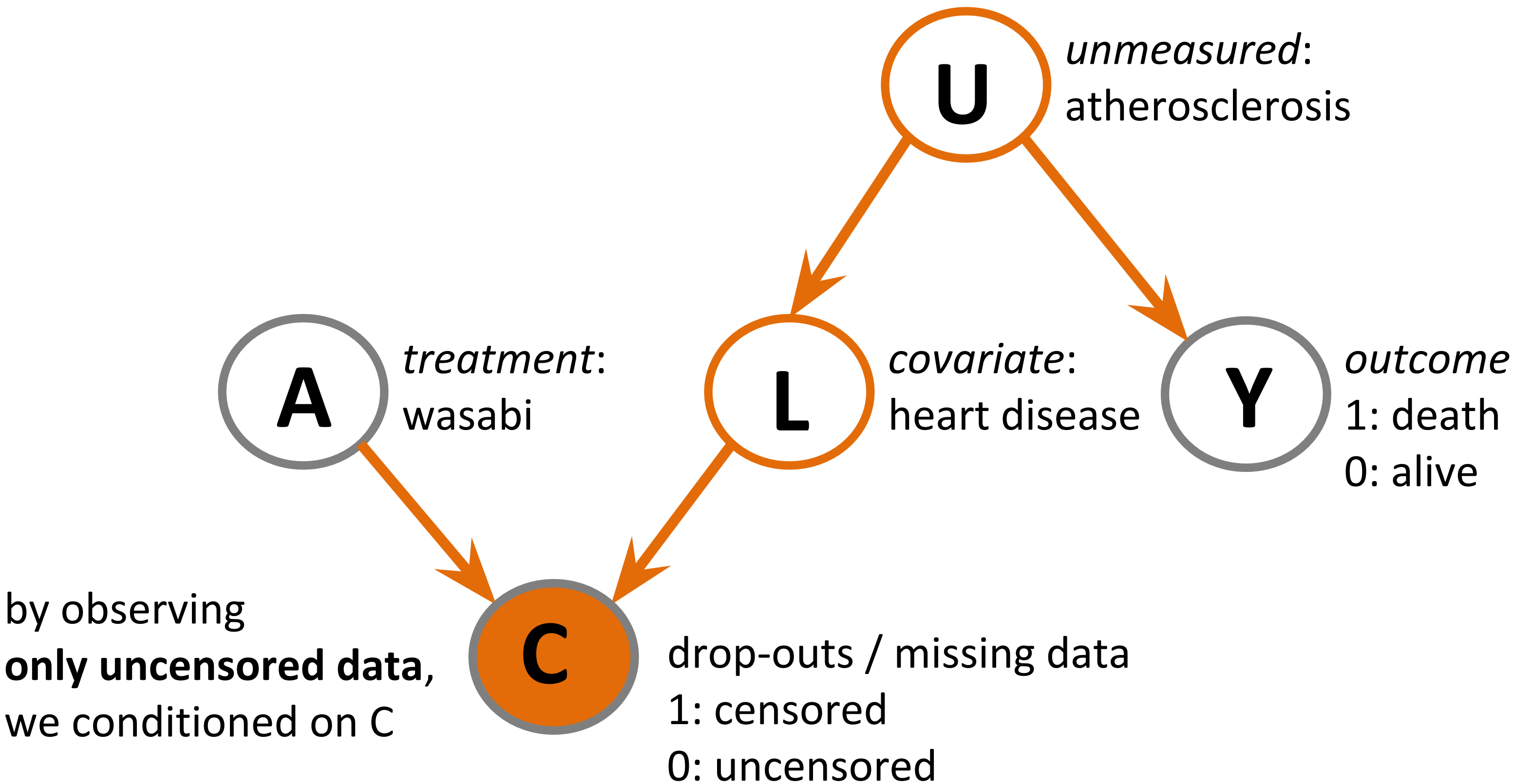
case II: we do **NOT** get data from all participants



Source: Causal Inference; Hernan and Robins; 2019

8.4 Selection Bias and **censoring**

case II: selection bias – **conditioning on collider (backdoor path)**

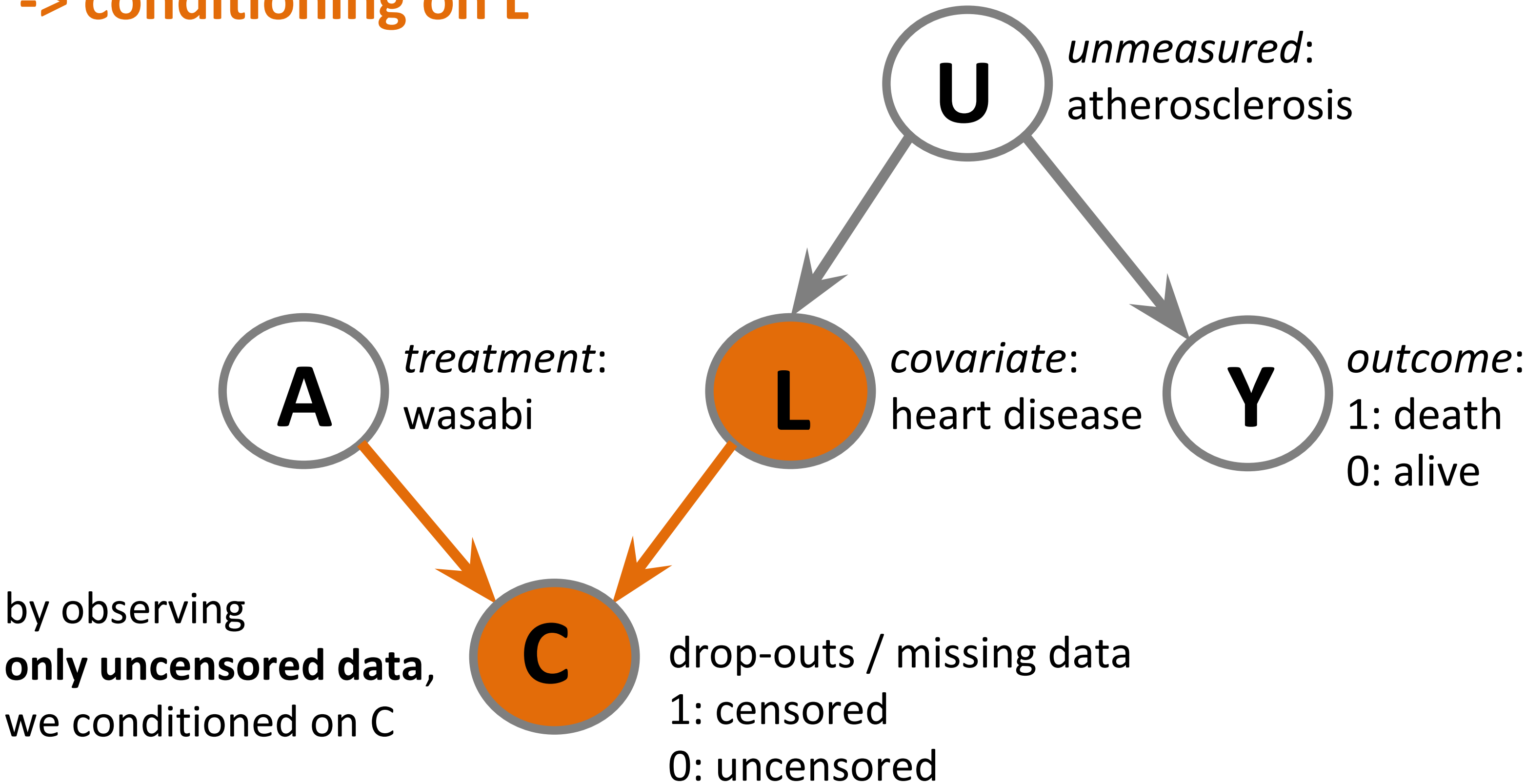


Source: Causal Inference; Hernan and Robins; 2019

conditioning is represented by colored node

8.4 Selection Bias and **censoring**

case III: selection bias – **blocking the backdoor path**
-> conditioning on L



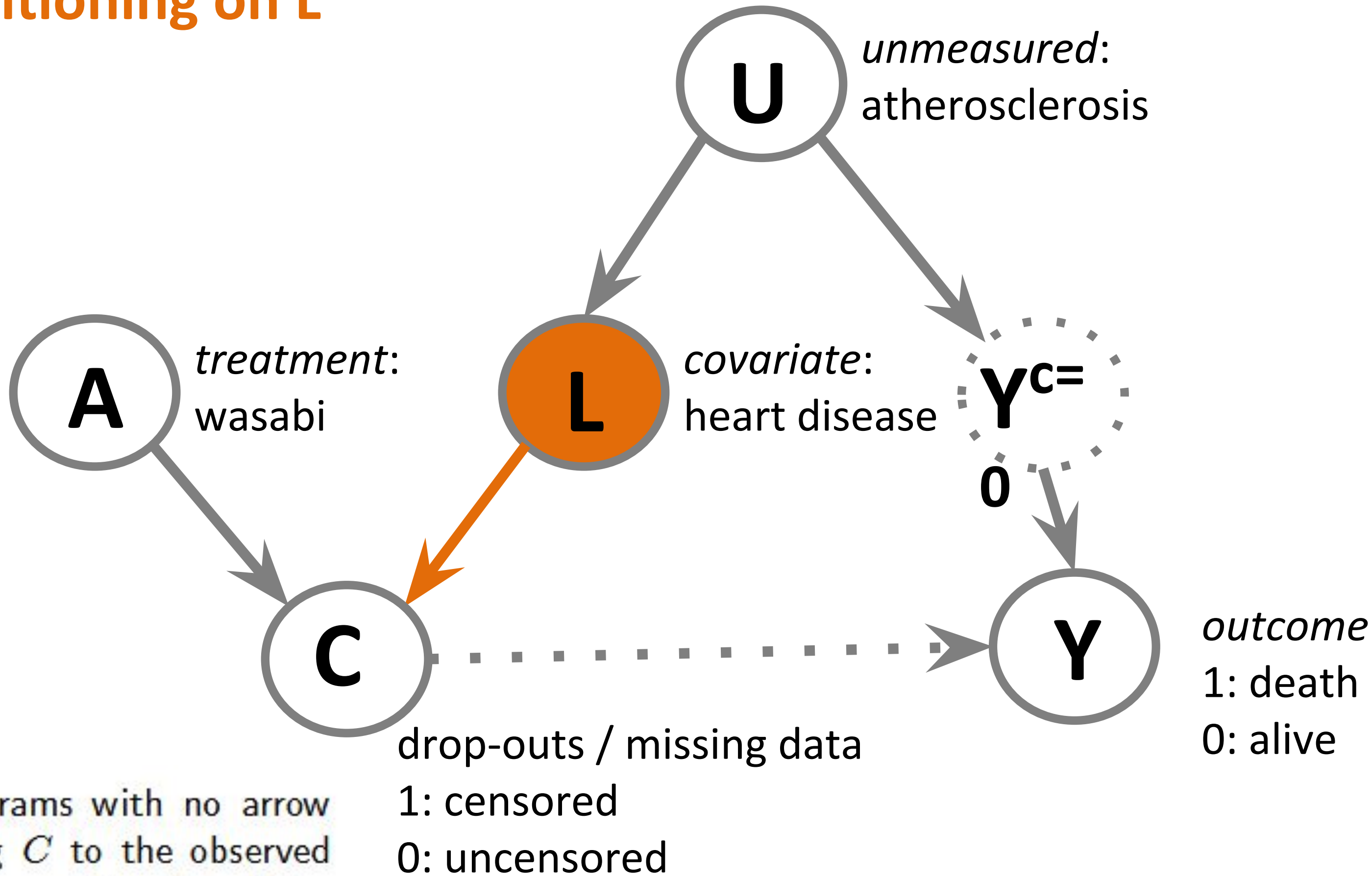
conditional risk ratio:
Pr[Y=1|A=1,C=0,L=l] / Pr[Y=1|A=0,C=0,L=l]

Source: Causal Inference; Hernan and Robins; 2019

conditioning is represented by colored node

8.4 Selection Bias and **censoring**

case III: selection bias – **blocking the backdoor path -> conditioning on L**



page 107:

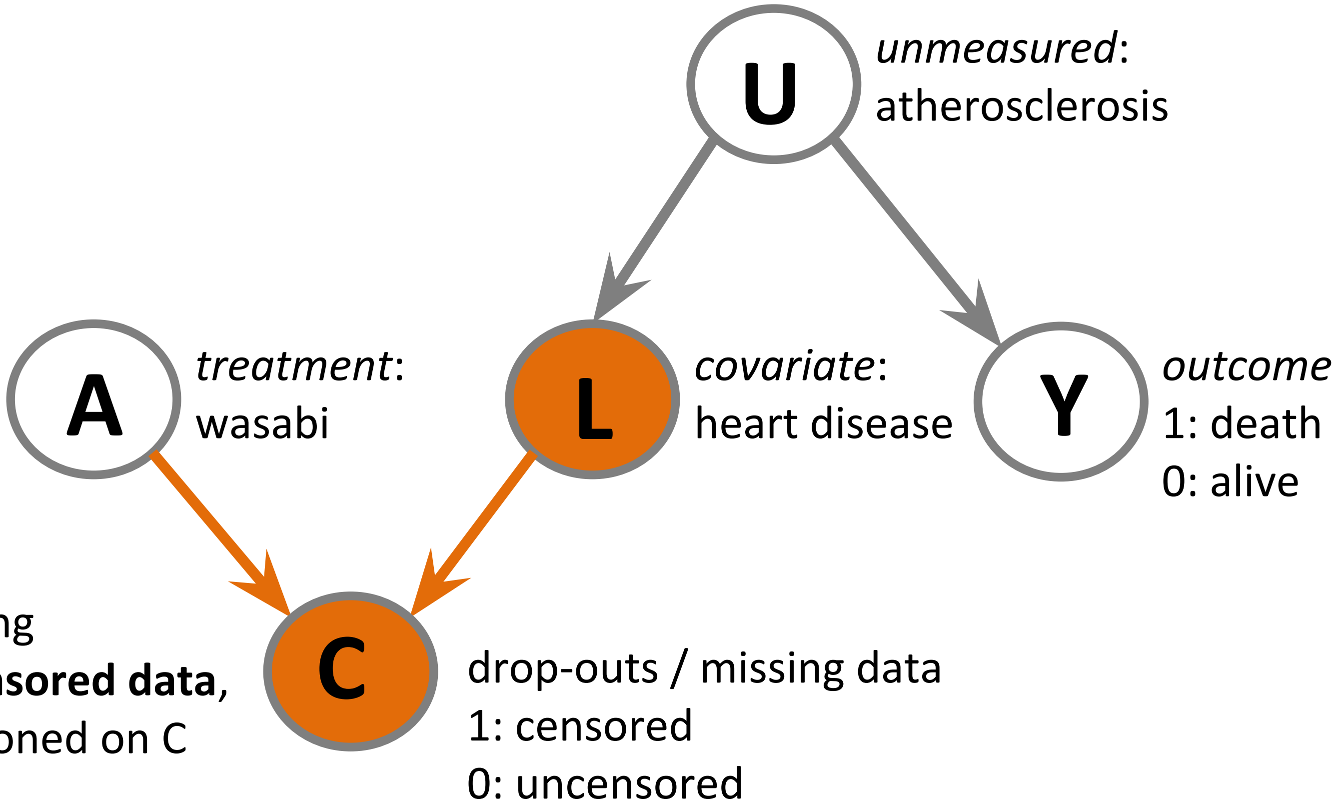
In causal diagrams with no arrow from censoring C to the observed outcome Y, we could replace Y by the counterfactual outcome $Y^{c=0}$ and add arrows $Y^{c=0} \longrightarrow Y$ and $C \longrightarrow Y$.

Source: Causal Inference; Hernan and Robins; 2019

conditioning is represented by colored node

8.5 How to **adjust** for selection bias

-> stratification



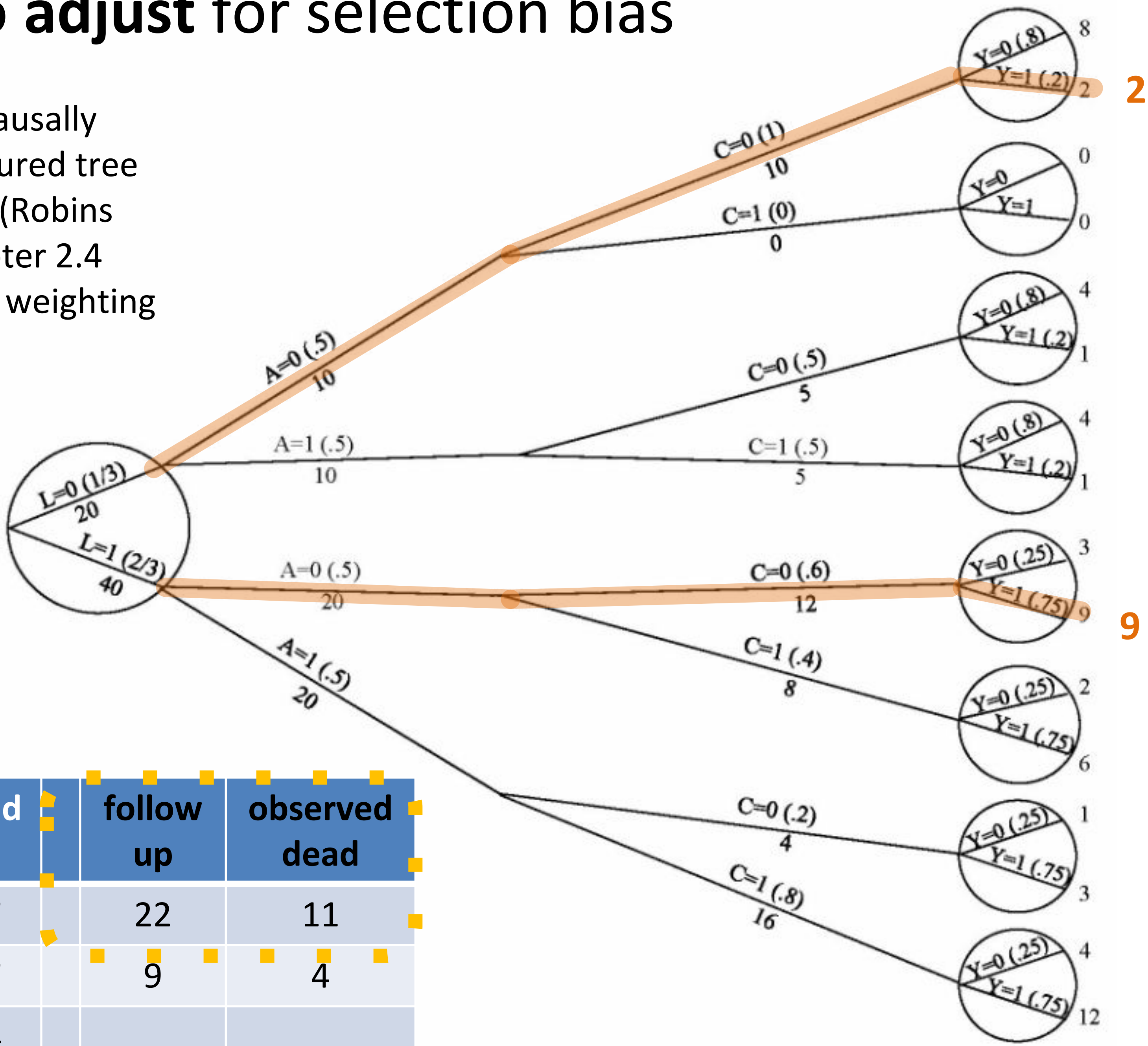
by observing
only uncensored data,
we conditioned on C

Source: Causal Inference; Hernan and Robins; 2019

conditioning is represented by colored node

8.5 How to **adjust** for selection bias

fully randomized causally inter-preted structured tree graph or **FR-CISTG** (Robins 1986, 1987) – chapter 2.4
Inverse probability weighting

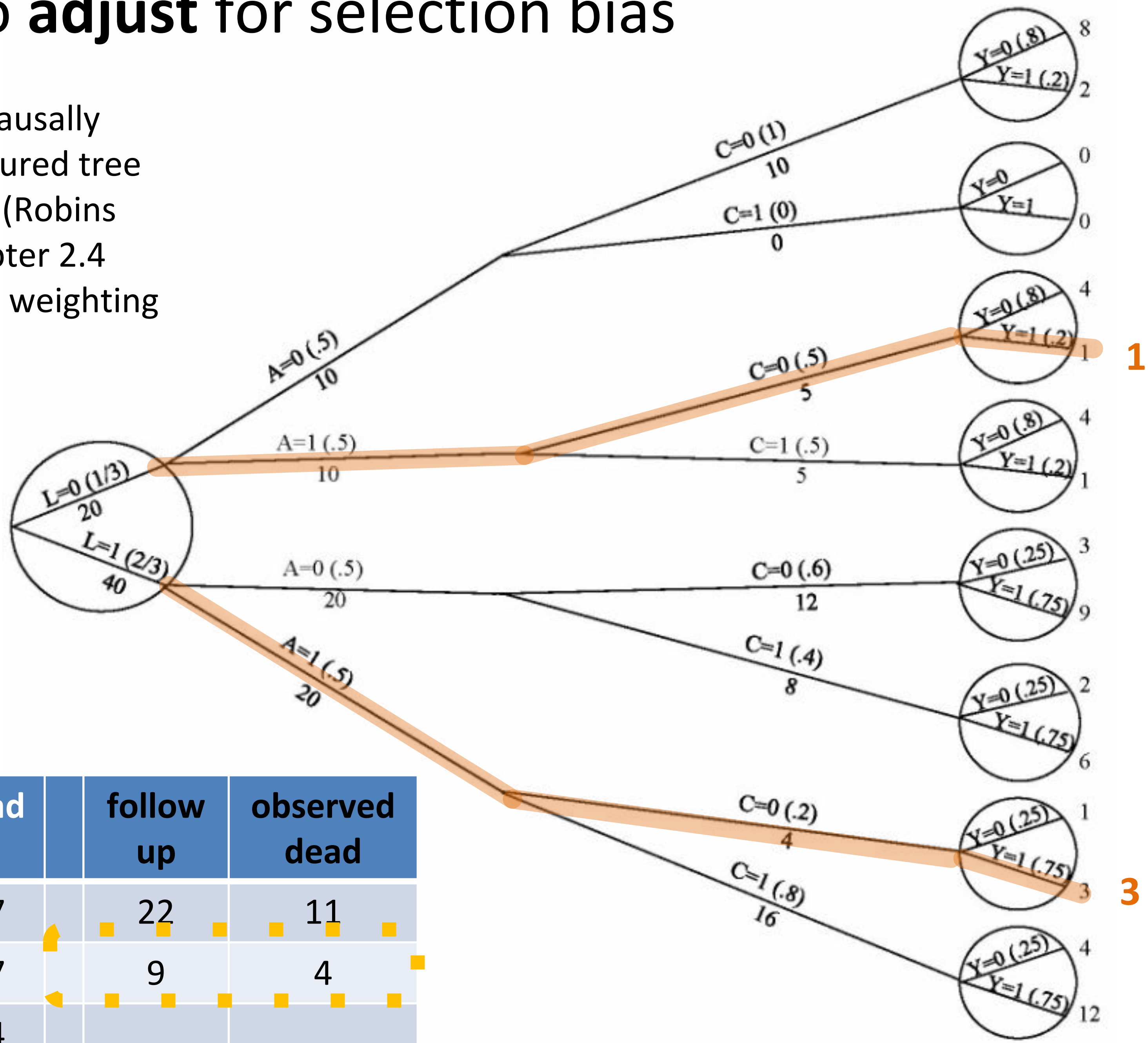


	alive	dead		follow up	observed dead
A=0	13	17		22	11
A=1	13	17		9	4
	26	34			

Source: Causal Inference; Hernan and Robins; 2019

8.5 How to **adjust** for selection bias

fully randomized causally inter-preted structured tree graph or **FR-CISTG** (Robins 1986, 1987) – chapter 2.4 Inverse probability weighting



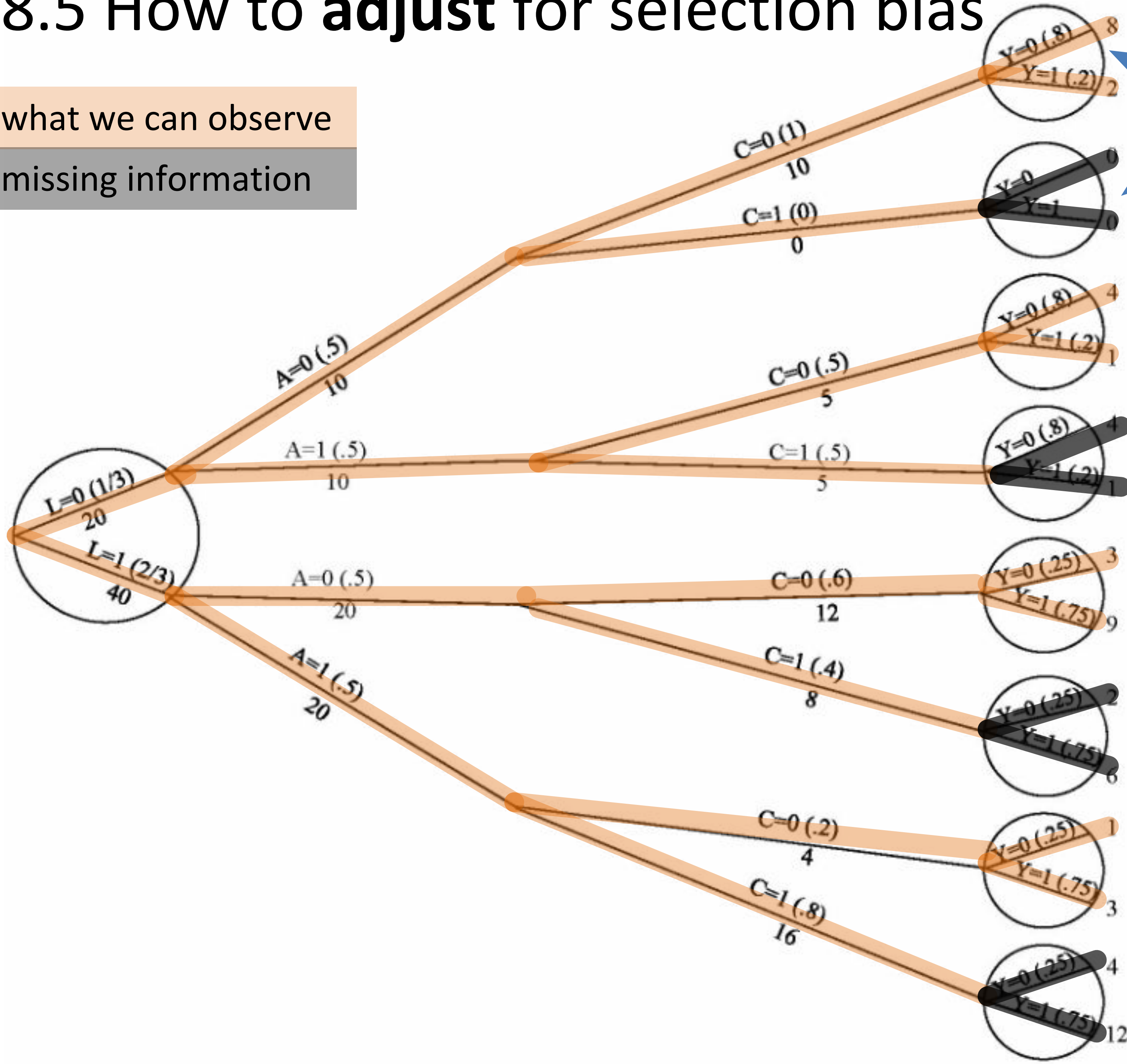
	alive	dead		follow up	observed dead
A=0	13	17		22	11
A=1	13	17		9	4
	26	34			

Source: Causal Inference; Hernan and Robins; 2019

8.5 How to **adjust** for selection bias

what we can observe

missing information



we assume:
conditional exchangeability given L,A

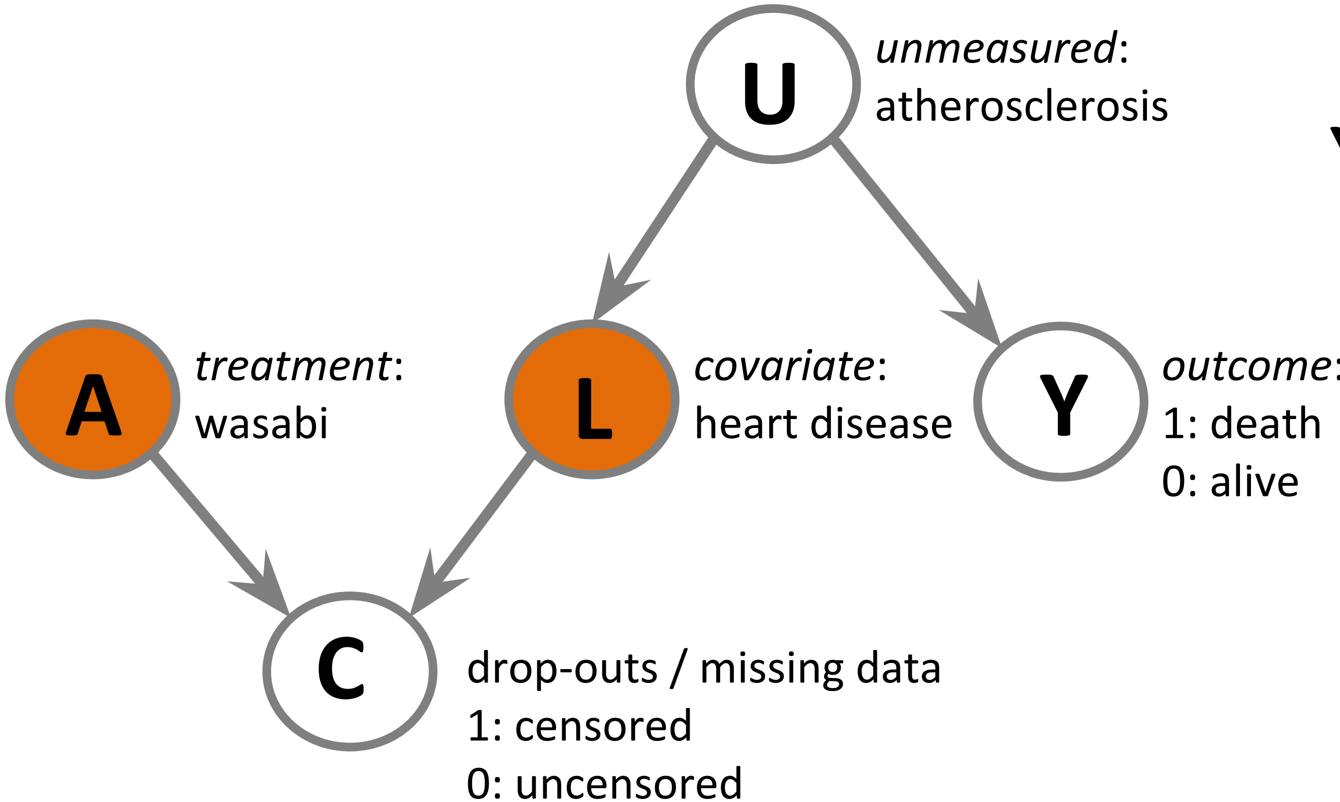
$$Y^C \perp\!\!\!\perp C$$

Source: Causal Inference; Hernan and Robins; 2019

8.5 How to **adjust** for selection bias

we assume:
conditional exchangeability given L,A

$Y^c \perp\!\!\!\perp C$



page 109:

conditioning on A and L is sufficient in blocking the backdoor path
C ← L ← U → **Y** between **C** and **Y**

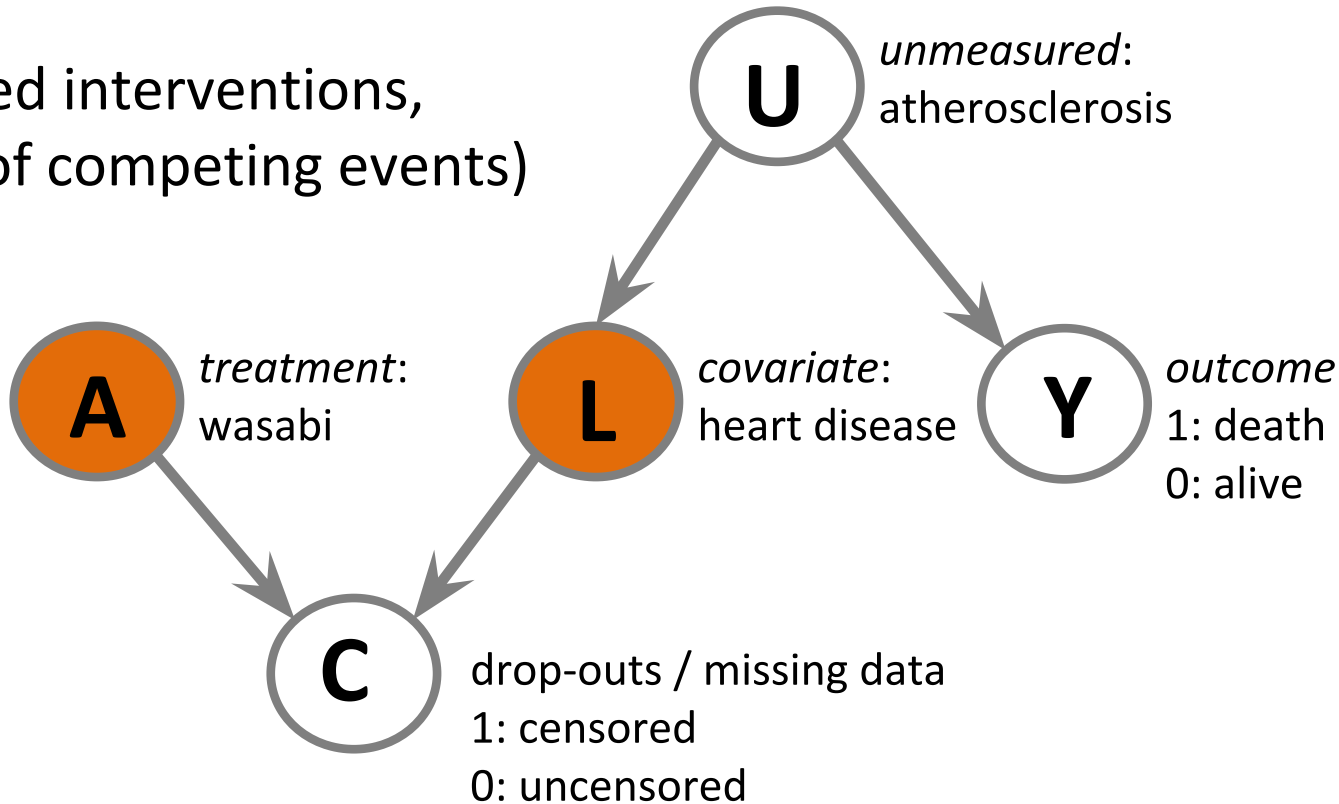
Source: Causal Inference; Hernan and Robins; 2019

conditioning is represented by colored node

8.5 How to **adjust** for selection bias

identifiability conditions:

- **exchangeability**
- **positivity**
- **consistency**
  (well defined interventions,
  be aware of competing events)



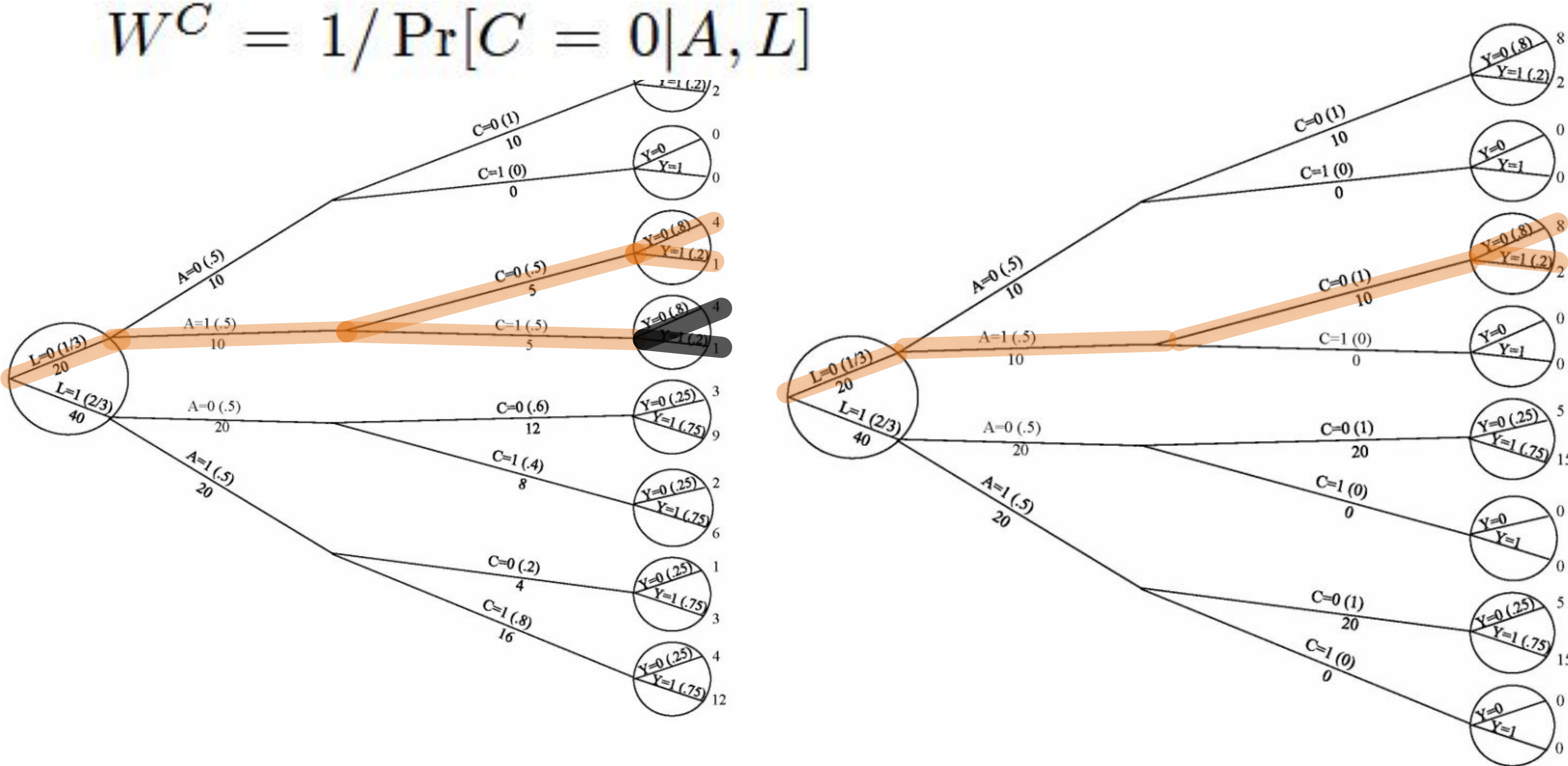
Source: Causal Inference; Hernan and Robins; 2019

conditioning is represented by colored node

8.5 How to **adjust** for selection bias

-> inverse probability (IP) weighting

$$W^C = 1/\Pr[C = 0|A, L]$$



use IP weight to redirect unobserved (C=1) population
to an estimated **pseudo-population**

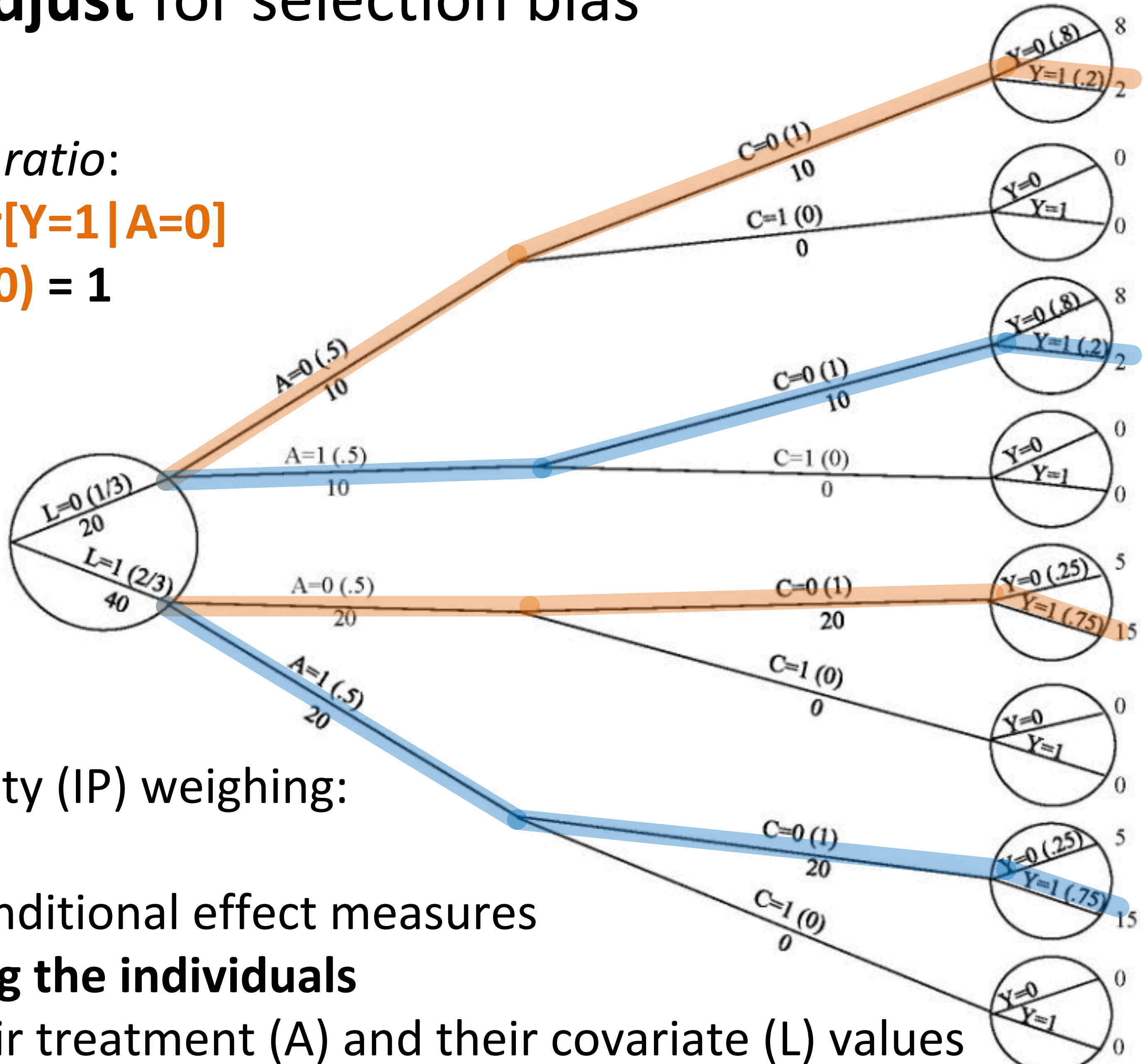
Source: Causal Inference; Hernan and Robins; 2019

8.5 How to **adjust** for selection bias

associational risk ratio:

Pr[Y=1|A=1] / Pr[Y=1|A=0]

= (17/30) / (17/30) = 1



inverse probability (IP) weighing:

estimating **un**conditional effect measures
after **reweighting the individuals**
according to their treatment (A) and their covariate (L) values

Source: Causal Inference; Hernan and Robins; 2019

8.5 How to **adjust** for selection bias

stratification

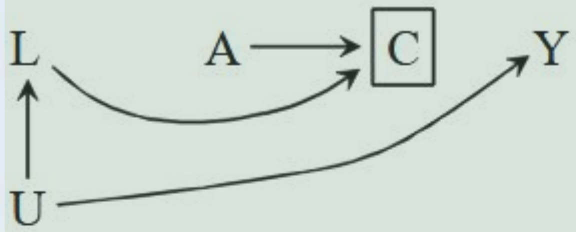


Figure 8.3

IP weighting

Source: Causal Inference; Hernan and Robins; 2019

8.5 How to **adjust** for selection bias

stratification

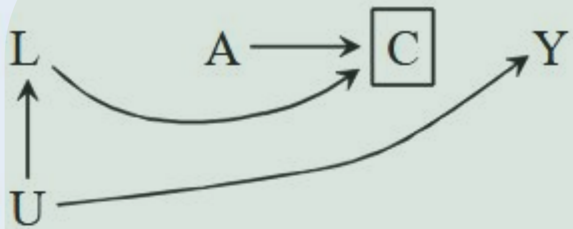


Figure 8.3

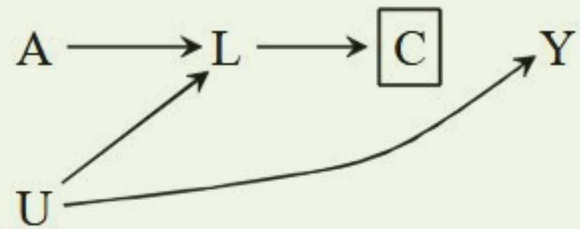


Figure 8.4

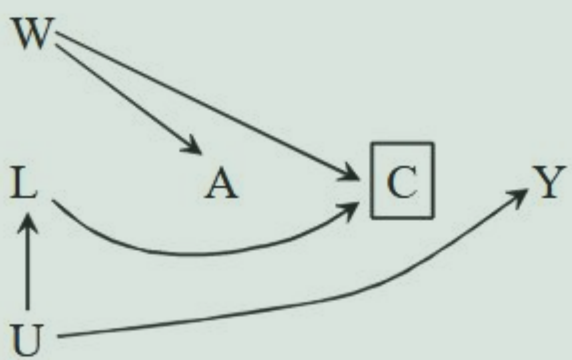


Figure 8.5

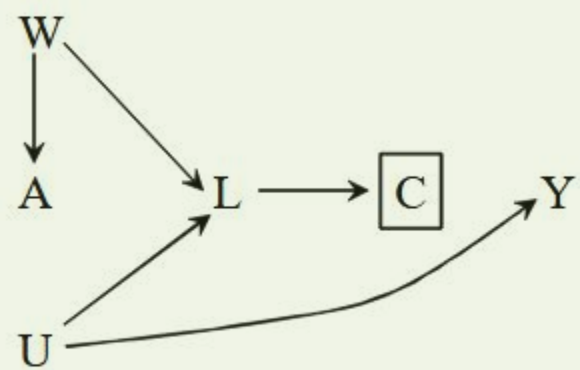


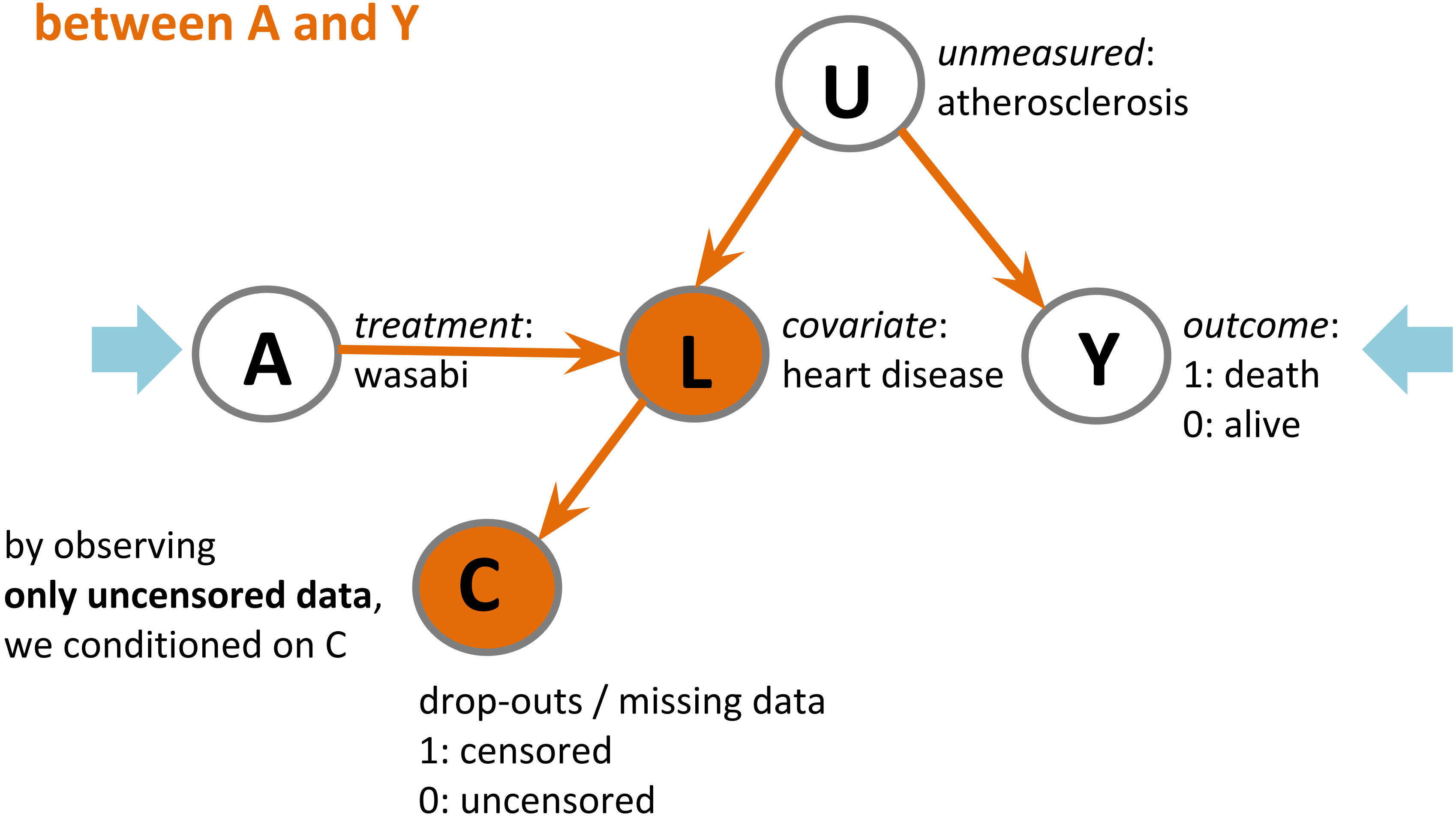
Figure 8.6

IP weighting

Source: Causal Inference; Hernan and Robins; 2019

8.4 Selection Bias and **censoring**

-> conditioning on L (stratification) opens backdoor path between A and Y



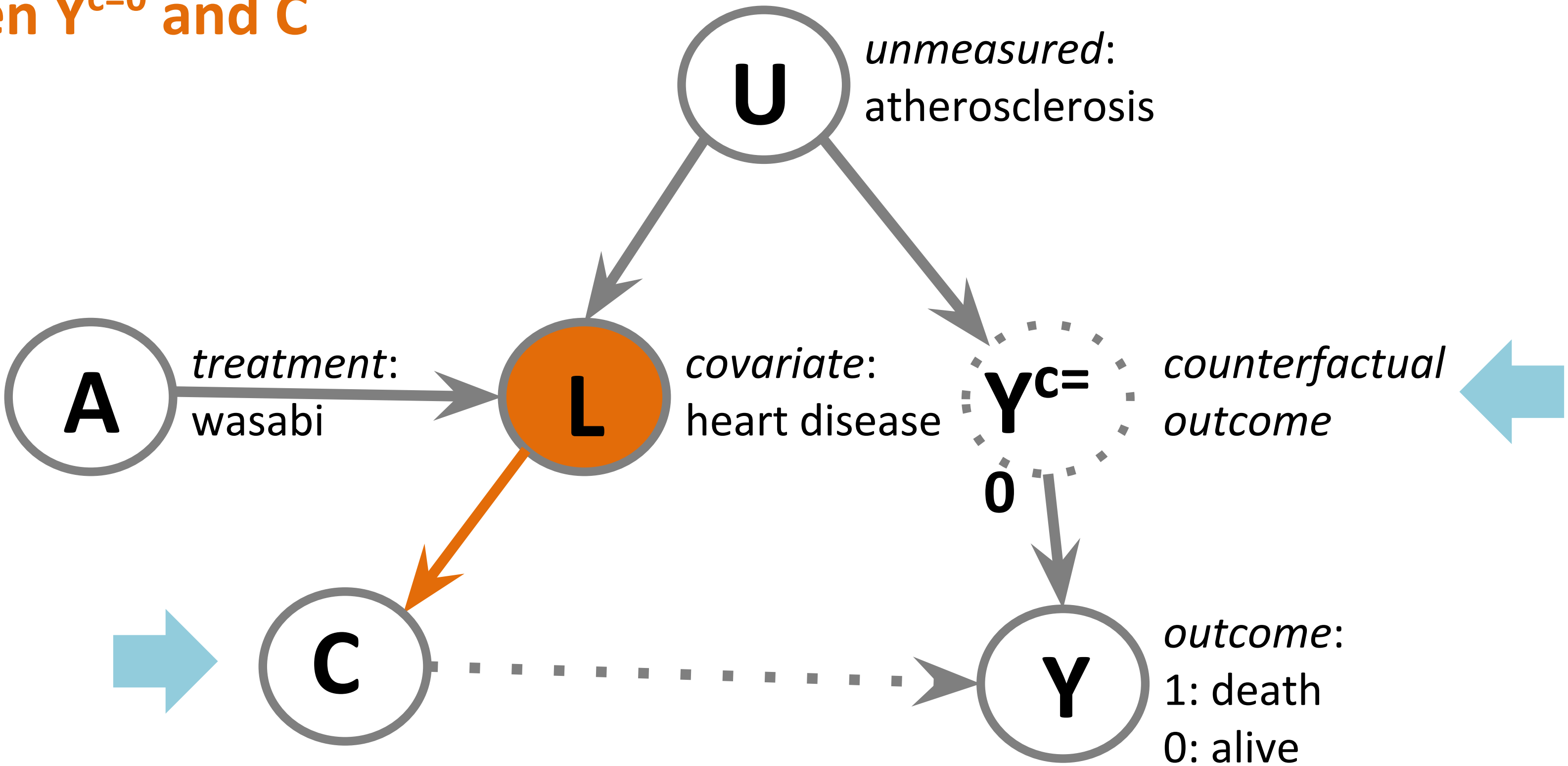
***our focus** in verifying exchangeability*

conditioning is represented by colored node

Source: Causal Inference; Hernan and Robins; 2019

8.4 Selection Bias and **censoring**

-> conditioning on L (IP weighting) d-seperates backdoor path between $Y^{c=0}$ and C



our focus *in verifying exchangeability*

conditioning is represented by colored node

Source: Causal Inference; Hernan and Robins; 2019

8.6 Selection without bias

conditioning on the common effect Y
induces conditional association between A and E

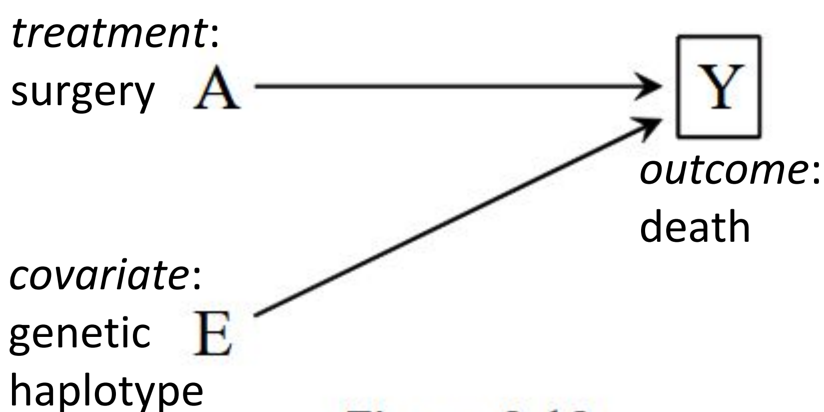


Figure 8.12

special situation (e.g. independent mechanisms):

dead Y=1: if $(Y_0=1)$ **or** $(Y_A=1)$ **or** $(Y_E=1)$
alive Y=0: if $(Y_0=0)$ **and** $(Y_A=0)$ **and** $(Y_E=0)$

Source: Causal Inference; Hernan and Robins; 2019

8.6 Selection without bias

augmented causal DAGs:
sufficient causes chapt. 5

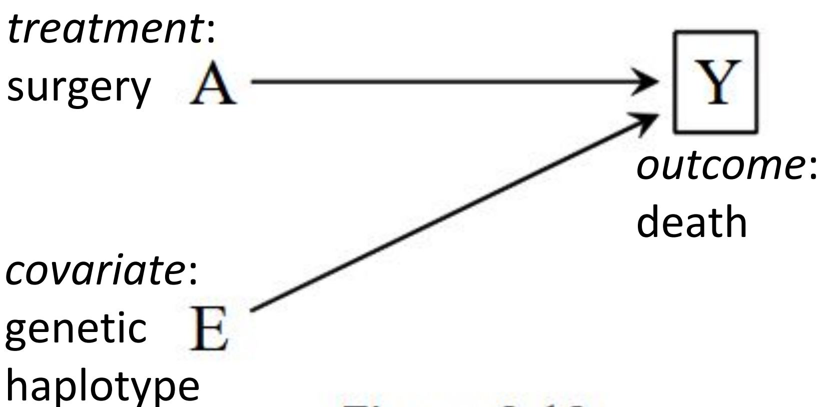


Figure 8.12

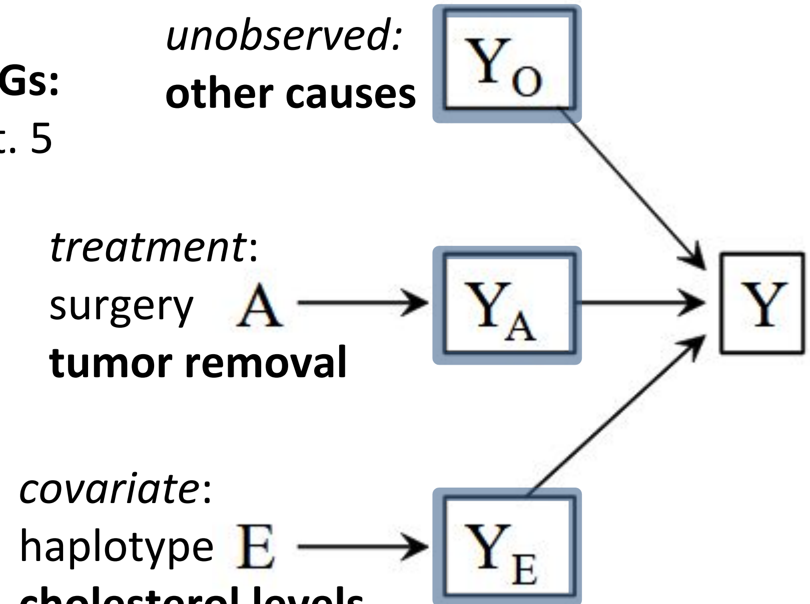


Figure 8.13

special situation (e.g. independent mechanisms):
A remains conditionally independant to E
even if conditioning on the common effect Y within one stratum

dead Y=1: if $(Y_0=1)$ **or** $(Y_A=1)$ **or** $(Y_E=1)$

enforced determinism: alive Y=0: if $(Y_0=0)$ **and** $(Y_A=0)$ **and** $(Y_E=0)$

Source: Causal Inference; Hernan and Robins; 2019

8.6 Selection without bias

the data in fig 8.13 follow
a multiplicative survival model

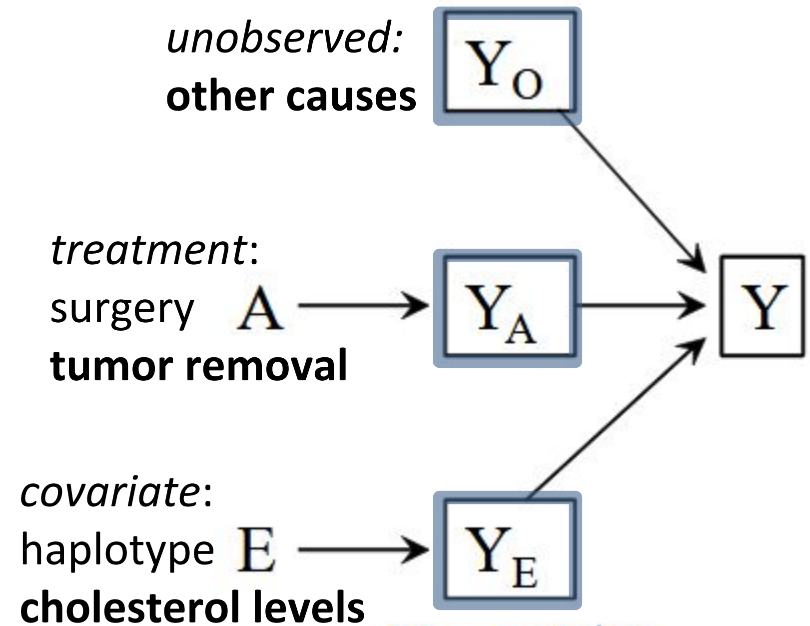


Figure 8.13

Technical Point 8.2

Multiplicative survival model. When the conditional probability of survival $\Pr[Y = 0|E = e, A = a]$ given A and E is equal to a product $g(e)h(a)$ of functions of e and a, we say that a multiplicative survival model holds. A multiplicative survival model

$$\Pr[Y = 0|E = e, A = a] = g(e)h(a)$$

is equivalent to a model that assumes the survival ratio $\Pr[Y = 0|E = e, A = a] / \Pr[Y = 0|E = e, A = 0]$ does not depend on e and is equal to $h(a)$. The data follow a multiplicative survival model when there is no interaction between A and E on the multiplicative scale as depicted in Figure 8.13. If $\Pr[Y = 0|E = e, A = a] = g(e)h(a)$, then $\Pr[Y = 1|E = e, A = a] = 1 - g(e)h(a)$ does not follow a multiplicative mortality model. Hence, when A and E are conditionally independent given $Y = 0$, they will be conditionally dependent given $Y = 1$.

Source: Causal Inference; Hernan and Robins; 2019

8.6 Selection without bias

unobserved:
other causes

treatment:
surgery
tumor removal

covariate:
haplotype
cholesterol levels

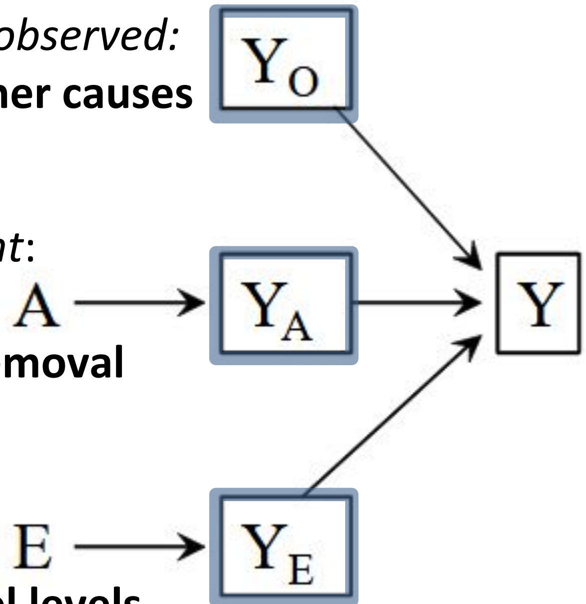


Figure 8.13

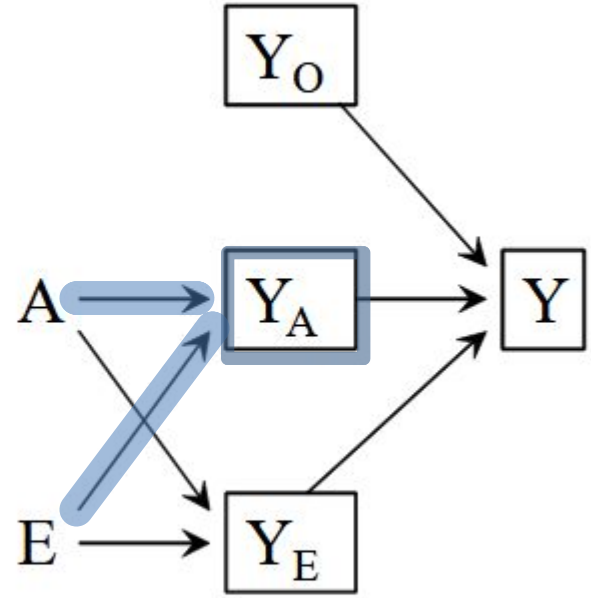


Figure 8.14

common mechanisms:
conditioning on common effects renders A and E conditionally dependant

interpretation:
conditioning on a collider always **induces an association** beetween its causes,
but this association **could be restricted to certain levels** of the common effect.
Collider stratification is not always a source of selection bias.

Source: Causal Inference; Hernan and Robins; 2019